Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

The development of a hyperspectral imaging method for the detection of *Fusarium*-damaged, yellow berry and vitreous Italian durum wheat kernels

*Silvia Serranti**, *Daniela Cesare*, *Giuseppe Bonifazi*

Department of Chemical Engineering Materials and Environment, Sapienza – University of Rome, Via Eudossiana, 18 00184 Roma, Italy

ARTICLE INFO

Article history:
Received 1 August 2012
Received in revised form
8 January 2013
Accepted 25 January 2013
Published online 17 March 2013

The possibility of using hyperspectral imaging (HSI) techniques to classify different types of wheat kernels, vitreous, yellow berry and *Fusarium*-damaged, was investigated. Conventional optical techniques adopted by industry for wheat grain sorting usually have too high misclassification errors. Reflectance spectra of selected wheat kernels of the three types were acquired by a laboratory device equipped with an HSI system working in near infrared field (1000–1700 nm). The hypercubes were analysed applying different chemometric techniques, such as principal component analysis (PCA) for explorative purposes, partial least squares discriminant analysis (PLS-DA) for classification of the three wheat types and interval PLS-DA (iPLS-DA) for the selection of a reduced set of effective wavelength intervals. The study demonstrated that good classification results were obtained not only considering the entire investigated wavelength range, but also selecting only three narrow intervals of four wavelengths (1209–1230 nm, 1489–1510 nm and 1601–1622 nm) out of 121. The procedures developed could be utilised at industrial level for quality control purposes or for the definition of innovative sorting logics for wheat kernels after an extensive classification campaign, both at laboratory and industrial level, applied to a large wheat sample sets.

1. Introduction

Wheat is one of the most important staple foods in the world, being used as raw material for breads, cakes, cookies, pastries, crackers and pasta products. Identification of different wheat types, at kernel level, is an important aspect for the food grain industry all over the world, involving both sorting and/or quality control strategies.

Quality of wheat grains is a complex phenomenon influenced by several factors, genetic and/or environmental. It is usually judged by evaluation of some parameters such as the grain vitreousness, protein content, gluten content, etc. (Sardana, 2000). The quality of both wheat kernels and related products may be considered from different points of view, depending on the purpose for which it is utilised.

This study focussed on the identification of three different durum wheat (*Triticum durum*) types (vitreous, yellow berry and *Fusarium*-damaged) posing a real industrial problem related to the correct identification of such classes using conventional optical techniques.

Vitreous kernels are hard, glassy, translucent and amber coloured, whereas non-vitreous kernels appear starchy and

* *Corresponding author*. Tel.: +39 0644585360; fax: +39 0644585618.
E-mail address: silvia.serranti@uniroma1.it (S. Serranti).
1537-5110/$ – see front matter © 2013 IAgrE. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biosystemseng.2013.01.011

opaque. The vitreousness of durum wheat is considered by the wheat industry as an important quality factor, improving cooking quality and pasta colour, and is associated with a coarser granulation and higher protein content (Dowell, 2000).

Yellow berry durum wheat kernels affect flour quality and, consequently, the quality of products like pasta. Yellow berry kernels are characterised by an undesirable yellowish and soft appearance. However, variable amounts of yellow berry kernels mixed with vitreous ones are commonly accepted by the market.

Fusarium species cause fungal diseases that affect small grain cereals such as wheat, barley, and rye (Shahin & Symons, 2011). In wheat, the fungus invades the spikelets, causing kernel damage in the form of shrivelling, loss of weight, and discolouration. The presence of *Fusarium*-damaged kernels produces not only a price penalty, but also a food safety concerns because of mycotoxins found in the grains infected by the fungus.

With reference to the above described three wheat types that are often mixed together in a batch, it is important to develop an industrial sorting process, mainly to eliminate *Fusarium*-damaged kernels but avoiding at the same time the loss of yellow berry kernels that are considered acceptable and are mixed in variable amounts with vitreous kernels for many products. In addition, where a top wheat quality is required, also yellow berry kernels can be selected from the batch.

The visual inspection methods, or human-knowledge-based methods, are subjective, sometimes inconsistent and slow. However, chemical methods are destructive, and time consuming (Mares, 1993; Neethirajan, Jayas, & White, 2007). Both these approaches are thus not suitable to be applied *for on-line* inspection. Optical methods are increasingly being investigated, and extensively applied, in order to set up accurate detection techniques for the identification of damaged wheat kernels. To achieve the goal of industrial-scale on-line inspection, multi- and hyper-spectral imaging based techniques analysis have been utilised.

There has been increased interest in machine-vision-based-technology to assess physical properties of grain (Sapirstein & Bushuk, 1989; Shatadal, Symons, & Dexter, 1998) promoted by many researchers who have developed combined multi-spectral image acquisition, processing, and analysis techniques with advanced classification algorithms, to detect grain kernel characteristics, such as colour, texture, and various types of damage (Bacci, Rapi, Colucci, & Novaro, 2002; Luo, Jayas, & Symons, 1999; Ruan et al., 2001; Zayas, Bechtel, Wilson, & Dempster, 1994). Different detection architectures and devices have been investigated. Symons, VanSchepdael, and Dexter (2003) developed a machine vision based system to classify durum wheat kernels according to the degree of vitreousness. Wang, Zhang, Dowell, and Pearson (2005) reached an overall correct classification of 94.83% for vitreous, non-vitreous and mottled (piebald) kernels using transmitted images. Shahin, Dorrian, and Symons (2005), through the processing of reflected and transmitted images for vitreous kernels, reached a correctness in classification of about 91% and 87% in the training and test set, respectively. Several imaging devices exist on the market (e.g. Acurum (www.acurum.dupont.com) and Cervitec (www.foss.it)) that can measure the vitreousness of durum wheat samples. They usually perform the analysis utilising a colour camera, performing the analysis on single grains in reflected light.

Other advanced methods, including X-ray imaging (Haff & Slaughter, 2004; Karunakaran, Jayas, & White, 2003) and near-infrared (NIR) spectroscopy, have shown potential for real-time applications. The X-ray based imaging approach is characterised by a high risk with reference to human health and for this reason, if in principle it is suitable for detection, it is practically not utilised. However, NIR spectroscopy has been more and more utilised for quality evaluation of many cereal grains (Singh, Paliwal, Jayas, & White, 2006) including detection of insect and insect parts in whole grain and ground samples (Baker, Dowell, & Throne, 1999; Dowell, Throne, & Baker, 1998; Dowell, Throne, Wang, & Baker, 1999; Maghirang, Dowell, Baker, & Throne, 2003). The NIR technique uses the spectral differences between healthy and not-healthy kernels caused by differences in chemical composition of healthy and damaged kernels for the discrimination.

The NIR spectroscopic based approach is affected by drawbacks such as imprecise estimates, the need to develop robust calibration models, preliminary processing of the collected data and, finally, the definition and implementation of robust classification logics. However, an imaging system working in the NIR shows several advantages, in respect of X-ray imaging, mainly in terms cost, simplicity, compactness and safety (Davies, 2000; Sugiyama, 1999).

Conventional NIR spectroscopic devices are usually point-based scanning instruments, providing only one spectrum of the target sample without giving any information about the distribution of the chemical composition of the sample. Spatial information is important for monitoring of the grain as it can be used to extract the chemical mapping of the sample starting from the collected spectra. Such a goal could be reached by utilising hyperspectral imaging (HSI) based devices.

The main problem faced in sorting processes based on optical sensors is the difficulty in the misclassification of yellow berry and *Fusarium*-damaged wheat kernels. Consequently, variable amounts of yellow berry kernels are rejected along with *Fusarium*-damaged ones, loosing in this way some quantities of good product with an economic loss.

Therefore, the development of an objective and non-destructive method for fast classification of the different wheat kernel types would be of benefit to producers, grain handlers, wheat millers and processors (Mahesh, Manickavasagan, Jayas, Paliwal, & White, 2008).

HSI techniques, as mentioned earlier, represent an attractive solution for characterisation, classification and quality control not only in the specific field of wheat grains, but also with reference to many different materials/products found in several industrial sectors. Recently, HSI has rapidly emerged and is rapidly developing, especially in food inspection with a large range of investigated products, such as fruits and vegetables, meat, fish, eggs and cereals (Del Fiore et al., 2010; Gowen, O'Donnell, Cullen, Downey, & Frias, 2007; Sun, 2010), in pharmaceutical sector (Fortunato de Carvalho Rocha, Post Sabin, Março, & Poppi, 2011; Gowen, O'Donnell, Cullen, & Bellc, 2008), in medicine (Blanco et al., 2012; Jolivot,

Vabres, & Marzani, 2011), in artworks (Kubik, 2007), in polymer science (Gosselin, Rodrigue, & Duchesne, 2011) and in the solid waste recycling sector (Bonifazi, Damiani, Serranti, Bakker, & Rem, 2009; Bonifazi & Serranti, 2006; Bonifazi, Serranti, Bonoli, & Dall'Ara, 2009; Leitner, Mairer, & Kercek, 2003; Serranti, Gargiulo, & Bonifazi, 2011, 2012; Tatzer, Wolf, & Panner, 2005).

HSI is based on the utilisation of integrated hardware and software architecture able to digitally capture and handle spectra, as an image sequence. In each image of the sequence, each column represents the discrete spectrum values of the corresponding element of the sensitive linear array. Such an architecture allows, with a "simple" arrangement of the detection device ("scan line" perpendicular to the moving direction of the objects) to realise a full and continuous control (Geladi, Grahn, & Burger, 2007; Hyvarinen, Herrala, & Dall'Ava, 1998). The spatial and spectral information, obtained simultaneously from the investigated object, are contained in a "hypercube", a 3D dataset characterised by two spatial dimensions and one spectral dimension. Considering that in a hyperspectral image the spectrum of each pixel can be analysed, HSI is the non-destructive technology providing the most accurate and detailed information extraction. According to the different wavelengths of the source and the spectral sensitivity of the device, several physical–chemical characteristics of a sample can be investigated and analysed.

Here, a HSI technique, working in the NIR range (1000–1700 nm), coupled with chemometric analysis, was developed and evaluated to classify different types of wheat, vitreous, yellow berry and *Fusarium*-damaged, in order to define a new method that can be implemented at industrial level for grain quality control and/or on-line sorting.

2. Materials and methods

2.1. Wheat samples

Selected samples of three different durum wheat kernel types were used: vitreous, yellow berry and *Fusarium*-damaged. An example of the analysed samples is shown in Fig. 1, in which the different visual features of the 3 types, according to the description reported in Section 1, can be appreciated.

The three sample sets are representative of the grain types that are mixed together in a wheat batch constituting the feed of an industrial sorting process. In this process, the *Fusarium*-damaged kernels must be eliminated and are therefore unacceptable in further treatments, such as grinding for flour production or for use in other food products. Yellow berry grains are acceptable with limits within vitreous grains, as previously outlined.

The output of the process was to be one or two good products (mixed or separated vitreous and yellow berry) and one waste product (*Fusarium*-damaged wheat), according to the final destination and use.

The three samples have been acquired by hyperspectral imaging according to different experimental set-up, that is:

Experimental set-up 1. There was bulk acquisition of each type. About 5 g of each sample were selected, corresponding to around 40 single grains. These sample sets have utilised to build the classification model.

Fig. 1 – The three types of studied durum wheat samples. 1a: Vitreous. 1b: *Fusarium*-damaged and 1c:yellow berry.

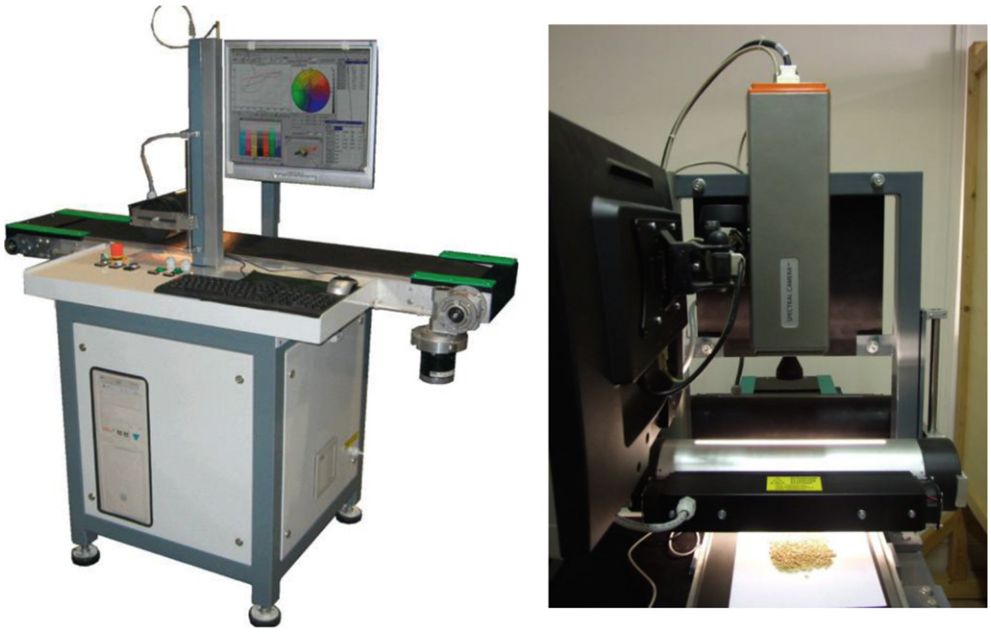
Experimental set-up 2. Three different and separated parallel lines, consisting of 20–30 single seeds each, of known type, but different from the seeds utilised to build the classification model, were acquired and processed. In the first line was placed the vitreous grains, in the second *Fusarium*-damaged grains and in the third the yellow berry grains. This second acquisition was carried out in order to validate the classification model developed through the acquisition and processing performed in the previous experimental set-up.

2.2. The hyperspectral imaging system

The HSI acquisition of the durum wheat samples was carried out at the Laboratory for Particles and Particulate Solids Characterisation (Latina, Italy) of the Department of Chemical Engineering, Materials & Environment ("Sapienza" University of Rome).

A specifically designed HSI-based platform (DV srl, Padova, Italy) was utilised to perform all the analyses (Fig. 2). The HSI based detection architecture allows not only static, but also dynamic analysis so it is possible to carry out tests on particle flow streams transported on a conveyor belt in order to perform, at laboratory scale, on-line particle detection from a sorting and/or quality control perspective.

The platform, in terms of hardware components, was based on a controlled conveyor belt (width = 260 mm and length = 1600 mm) with adjustable speed (variable between 0 and 50 mm s^{-1}) (Fig. 2). Spectra acquisition was carried out continuously or at specific time intervals. The utilised acquisition system was an NIR Spectral Camera™ (Specim, Oulu, Finland), containing an ImSpector N17E™ imaging spectrograph working in spectral range from 1000 to 1700 nm, with a spectral sampling/pixel of 2.6 nm, coupled with a Te-cooled InGaAs photodiode array sensor (320 × 240 pixels) with pixel resolution of 12 bits. A diffused light cylinder source, providing the required energy for the sensing unit, was set up. The cylinder, internally coated with aluminium, contained



2a 2b

Fig. 2 – The hyperspectral imaging (HSI) platform – 2a): overview of the integrated HSI based architecture, 2b): detail showing the spectral camera NIR (Specim, Oulu, Finland) and the corresponding lighting sources.

five halogen lamps producing a continuous spectrum signal optimised for spectra acquisition in the NIR wavelength range.

The device worked as a push-broom type of line scan camera allowing the acquisition of spectral information for each pixel in the line (Hyvarinen et al., 1998). The transmission diffraction grating and optics provide high light throughput and high quality and distortion-less image for the device. The result of the acquisition was a digital image where each column represented the discrete spectrum values of the corresponding element of the sensitive linear array.

The device was controlled by a PC unit equipped with the Spectral Scanner™ v.2.3 (DV srl, Padova, Italy) acquisition/preprocessing software, specifically developed to handle the different units and the sensing device constituting the platform and to perform the acquisition and the collection of spectra. The software was designed with flexible architecture so it could be easily integrated with new software modules containing new characterisation and/or classification tools.

2.3. *Image acquisition and calibration*

Hyperspectral images of durum wheat were acquired in the 1000–1700 nm wavelength range, with a spectral resolution of 7 nm, for a total of 121 wavelengths. The spectrometer was coupled to a 15 mm lens. The images were acquired scanning the investigated sample line by line. The image width was 320 pixels, while the number of frames was 280.

Calibration was performed recording two images for black and white references. The black image (B) was acquired to remove the effect of dark current of the camera sensor, turning off the light source and covering the camera lens with its cap. The white reference image (W) was acquired adopting an NPL Spectralon (ProLite Technology, Innovation Centre, Cranfield University, Milton Keynes, London, UK) reference sample under the same condition of the raw image. Image correction was performed adopting the following:

$$I = \frac{I_0 - B}{W - B} \times 100 \quad (1)$$

where I is the corrected hyperspectral image in a unit of relative reflectance (%), I_0 is the original hyperspectral image, B is the black reference image (~0% reflectance) and W is the white reference image (~99.9% reflectance). All the corrected images were then used to perform the HSI based analysis, that is to extract spectral information, to select the effective wavelengths and for the final classification purposes.

3. Spectral data analysis

Spectral data analysis was carried out adopting standard chemometric methods (Geladi et al., 2007; Otto, 1999) utilising the PLS_Toolbox (Version 6.5.1, Eigenvector Research, Inc., Wenatchee, USA) running inside Matlab® (Version 7.11.1, The Mathworks, Inc., Natick, Massachusetts, USA).

3.1. *Spectra preprocessing and principal component analysis*

Firstly, the raw spectra were cut at the beginning and at the end of the wavelength range in order to eliminate unwanted effects due to background noise. The number of wavelengths was thus reduced from 121 to 92 and the new investigated interval was 1013–1650 nm.

Data were then pre-processed in order to highlight the differences between the three wheat types. The generalised least squares weighting (GLSW) algorithm was applied. It calculates

a filter matrix based on the differences between pairs or groups of samples which should otherwise be similar (Martens, Høy, Wise, Bro, & Brockhoff, 2003). Considering that hyperspectral cameras produce a huge amount of data (the hypercubes acquired in this study have a size of 320 × 280 × 92), chemometric processing is needed for data exploration and modelling. Therefore, after pre-processing, an exploratory analysis was carried out applying principal component analysis (PCA) to the spectral data (Wold, Esbensen, & Geladi, 1987). PCA compresses the data by projecting the samples into a low dimensional subspace, whose axes (the principal components, PCs) point in the directions of maximal variance. Looking at the distribution of the samples in PC space it is possible to analyse their common features and/or their grouping.

The first few principal components (PCs), resulting from PCA, are generally utilised to analyse the common features among samples and their grouping: samples characterised by similar spectral signatures tend, in fact, to aggregate in the score plot of the first two or three components. Therefore spectra could be characterised either by the reflectance at each wavelength in the wavelength space, or by their score on each PC in PC space.

Samples characterised by similar spectra, that is belonging to the same class of products (in this case wheat), are thus grouped in the same region of the score plot related to the first two or three PCs, whereas samples with different spectral features will be clustered in other parts of this space.

3.2. *Partial least square-discriminant analysis*

PCA is a powerful method for data exploration, being able to highlight the presence of trends or clusters among samples. However, PCA is an unsupervised technique and it cannot be used for building predictive model, for instance to classify samples in one or another category: in the latter case, a supervised pattern recognition approach should be adopted.

Partial least square discriminant analysis (PLS-DA) was used to find a model able to perform an optimal discrimination among classes of samples and for prediction in new images. PLS-DA is a supervised classification technique, requiring a prior knowledge of the data (Barker & Rayens, 2003); it is used to classify samples into predefined groups by forming discriminant functions from input variables (wavelengths) to yield a new set of transformed values that provides a more accurate discrimination than any variable (wavelength) alone. A discriminant function is then built using samples with known groups to be used later to classify samples with unknown group membership. Therefore, once the model is obtained, it can be applied to an entire hypercube, and for classification of new hypercubes. The result of PLS-DA applied to hyperspectral images is a "prediction map", where the class of each pixel can be identified using colour mapping. The optimal dimensionality of the PLS-DA classification model was defined using contiguous-blocks cross-validation (10 deletion groups). The purpose of PLS-DA applied to wheat samples was to validate their correct classification using both all the wavelengths and the effective wavelengths selected applying the interval PLS-DA (iPLS-DA) method (see Section 3.3). The hyperspectral image of wheat in bulk was used as training set, while the image wheat kernels in parallel lines was used as test set to quantitatively evaluate the predictive ability of the calculated classification models.

3.3. *Wavelength selection by interval PLS-DA*

Because the extracted spectral data from wheat images are characterised by a high degree of dimensionality with redundancy among contiguous variables (wavelengths), a selection of wavelengths was carried out, in order to facilitate and speed up the classification of the three different wheat classes. In this study the iPLS-DA variable selection method was adopted (Nørgaard et al., 2000). iPLS-DA does a sequential, exhaustive search for the best variable or combination of variables. Furthermore, it can be operated in "forward" mode, where intervals are successively included in the analysis, or in "reverse" mode, where intervals are successively removed from the analysis.

The "interval" in iPLS can be either a single variable or a "window" of adjacent variables. iPLS-DA works by dividing the full-spectrum in intervals of equal width and calculating classification models for each one of these spectral regions. In the forward-selection mode, the best interval is then chosen as the one leading to the minimum value of the root mean square error in cross-validation (RMSECV). Then, two-interval models are built by adding each one of the remainder intervals to the previously selected one. Once again, the model showing the lowest RMSECV value is selected and this iterative procedure is repeated until no significant improvement of RMSECV is achieved. Conversely, in the reverse-selection mode, the intervals are iteratively removed according to a decrease in the RMSECV value. Since this latter procedure is more conservative than the former one, i.e. a larger number of wavelengths is usually preserved in the final model, the forward-selection mode was used in the present work.

In particular, forward iPLS-DA was applied to the training image with wheat in bulk considering a window size of 4 variables (23 intervals). The same preprocessing and cross-validation procedures used for the calculation of the PLS-DA models on the whole signals were applied.

4. Experimental results

4.1. *Spectral features of the different durum wheat kernel types*

The investigated NIR range provides chemical information about cereal composition because most absorption bands observed in this region arise from overtones of C–H, O–H and N–H stretching vibrations (Moller & Munch, 2009). The average raw NIR reflectance spectra of the three different classes of durum wheat are reported in Fig. 3. The trough at 1200 nm is related to the C–H second overtone, the main one at 1450 nm is linked to the O–H and N–H first overtone and to the C–H combination band in the aromatic ring. In Fig. 4 the preprocessed spectra obtained after the application of the GLSW algorithm are reported. Looking at the modified spectra, it can be seen that the differences between the three wheat types are highlighted after the application of the GLSW preprocessing.

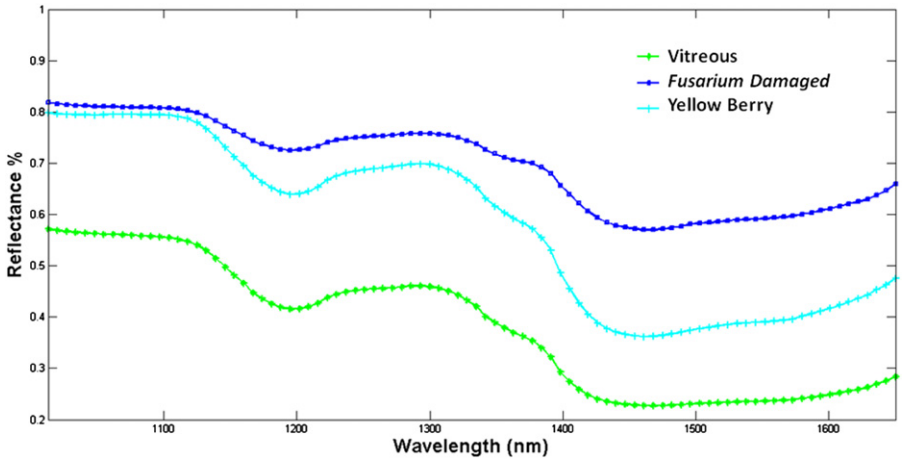


Fig. 3 – Comparison between the raw NIR average reflectance spectra of the three classes of durum wheat: vitreous, *Fusarium*-damaged and yellow berry.

4.2. *Explorative analysis and class definition*

From the hyperspectral image of wheat, in bulk, several regions of interest (ROIs) on wheat kernels, representative of vitreous (238 pixels), *Fusarium*-damaged (192 pixels) and yellow berry kernels (200 pixels) respectively, were selected (Fig. 5) and the corresponding spectra were arranged in a two dimensional matrix which constituted the training set for this investigation. Starting from the training matrix, made up of the spectra from the selected ROIs, exploratory data analysis was performed. The training data, pretreated as previously described applying the GLSW algorithm, were then used to build a PCA model for exploratory purposes.

The results obtained from the application of PCA to the three classes of wheat kernels are reported in Fig. 6, showing the score plot of PC1 *vs* PC2. The first two PCs explained 73.88% (53.19% and 20.69%, respectively) of variance. The spectral data of the three classes were clustered into three distinct groups according to their spectral signatures. The procedure allows a clear discrimination to be seen between the classes of wheat samples. In particular, it is evident that vitreous kernels were well separated from the other two classes along the first principal component, whereas *Fusarium*-damaged and yellow berry kernels were better distinguished along PC2.

Based on these results, PCA was applied to the entire hyperspectral image, considering that pixels with similar spectral patterns tend to be plotted in close proximity in PCA space they appear in similar colours in the classified/mapped image. Score images, resulting from the PCA model, showed in fact a different ability to visualise the main components of the examined samples based on the amount of captured variance. Since PCA performs a significant data compression, the first score images, obtained combining the first PCs components, allow to enhance the higher contrasts associated to the spectral composition of the image features. Score images, created utilising the first PCs contain, in fact, most of the variance of the original dataset, while the next PCs score image show less details and so on.

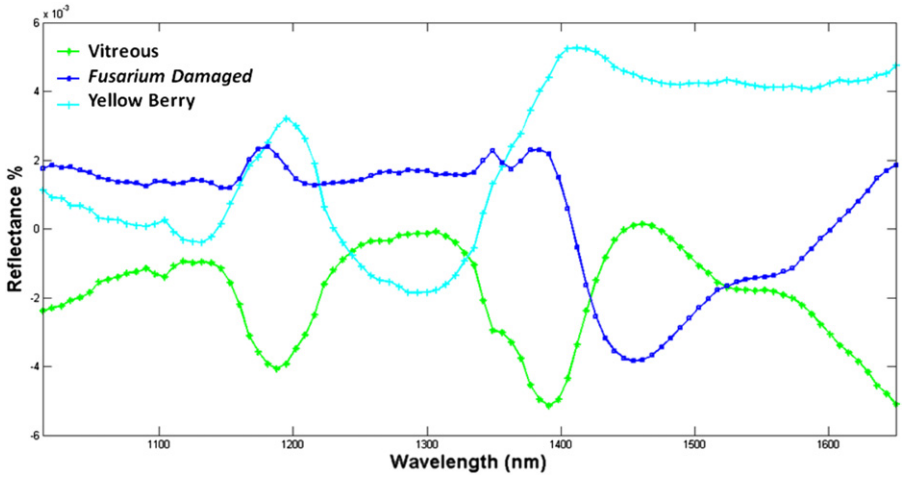


Fig. 4 – Comparison between the pre-processed NIR average reflectance spectra of the three classes of durum wheat: vitreous, *Fusarium*-damaged and yellow berry.

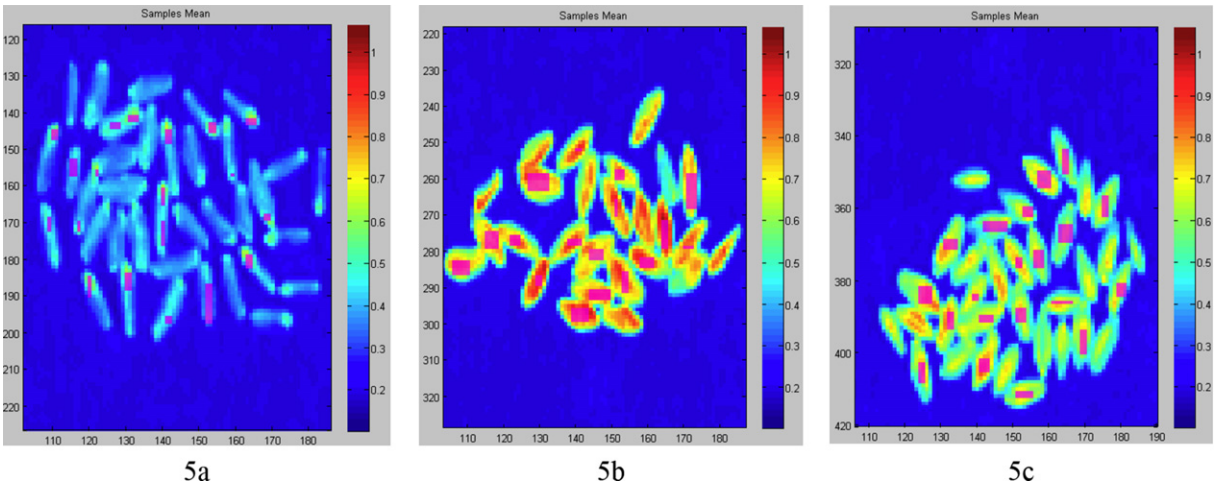


Fig. 5 – ROI selection (pink-coloured) on the spectral images of vitreous (5a), *Fusarium*-damaged (5b) and yellow berry (5c) durum wheat samples.

In this study the first two PC score images were combined together to form a pseudo-colour image (one score image representing one colour channel). In fact, as was discussed earlier, considering that the separation among the pixels corresponding to the three classes to discriminate occurs along the first and the second PCs, the corresponding score image was further investigated. These results are shown in Fig. 7 where the image background was masked and appears white. With the only exception of some boundary effects, linked to background removal, it can be seen that pixel colouring reflects perfectly the differences between the three classes, as already evidenced by the clear separation in Fig. 6.

4.3. *Wheat samples classification by partial least squares discriminant analysis*

A PLS-DA model was developed to discriminate between the three classes of wheat and for prediction in new images. The PLS-DA model was calculated using the same pre-treatment technique adopted before application of PCA (GLSW) and the same full cross validation method.

The PLS-DA model was then used to test the recognition/classification of the three wheat classes arranged in bulk (experimental set-up 1) (Fig. 8) and for the three wheat classes arranged in three parallel lines of single seeds (experimental set-up 2) (Fig. 9), being each line respectively constituted by vitreous (36 kernels), *Fusarium*-damaged (25 kernels) and yellow berry (32 kernels).

The results were visualised in Figs. 8a and 9b, in the form of a prediction image, i.e. of an image having the same dimension as the original one but where pixels are coloured according to the predicted category. Also in this case, white pixels corresponded to the background, which was masked. A correction was obtained in both cases for all the 3 classes, the only errors in prediction being related to pixels corresponding to the boundaries of the kernels, as shown in Figs. 8a and 9b.

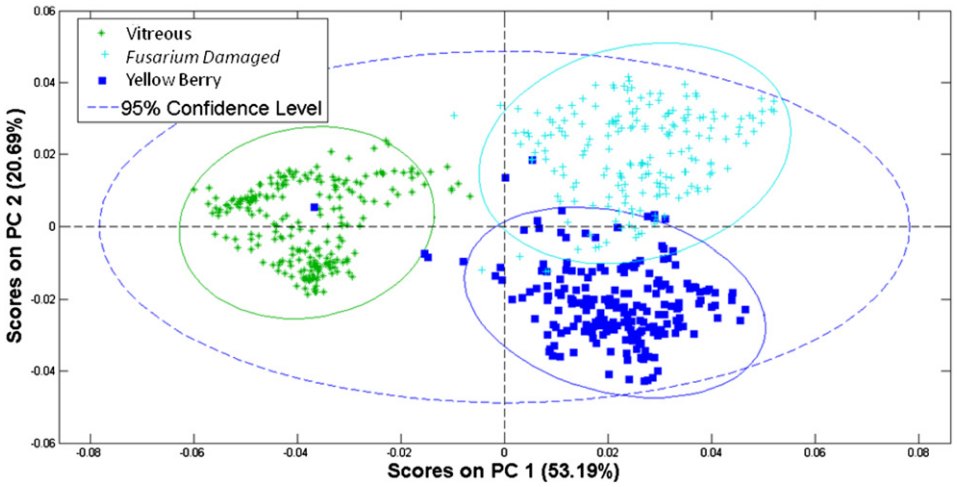


Fig. 6 – PCA score plot PC1 *vs* PC2 using 93 wavelengths in the whole spectral range (1006–1650 nm) related to the three durum wheat grain classes: vitreous, *Fusarium*-damaged and yellow berry.

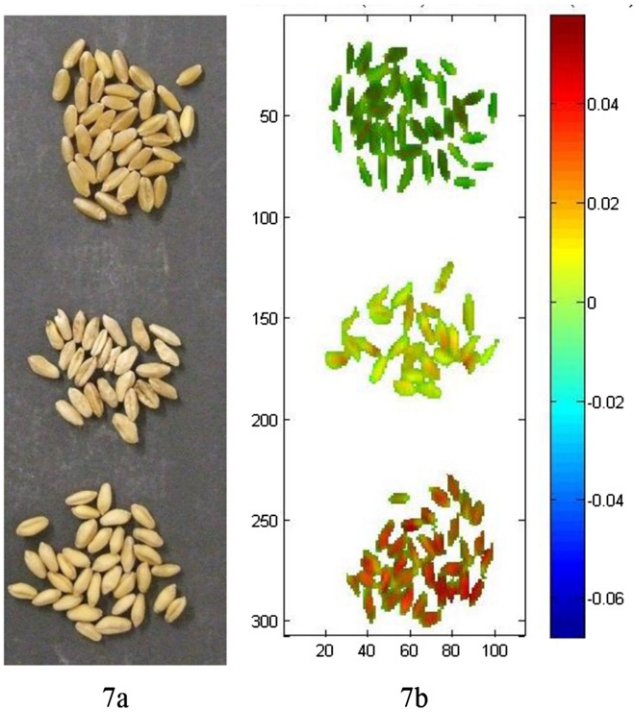


Fig. 7 – Comparison between digital image (7a) and classification image (7b) based on the PCA applied to the entire hypercube, as it results from the combination of the first two PCs score images. From top to bottom: vitreous, *Fusarium*-damaged and yellow berry kernels.

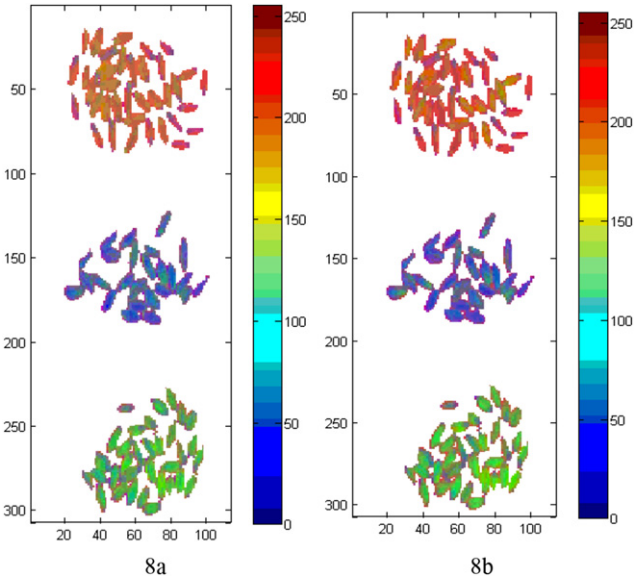


Fig. 8 – Prediction images based on PLS-DA model built for the classification of wheat in bulk (see Fig. 7a) using 92 wavelengths (8a) and 12 wavelengths, respectively (8b) (red: vitreous, blue: *Fusarium*-damaged and green: yellow berry).

4.4. *Wheat samples classification after variable selection by interval PLS-DA*

Having in mind the objective of using the proposed method for on-line process monitoring and/or control, the possibility of building classification models on a reduced number of wavelengths appears promising as it could decrease the time required to acquire and process each spectral image, opening the possibility of interesting scenarios for the design, the development and the set-up of innovative HSI based sorting procedures of wheat grains for quality control and/or separation purposes.

Accordingly, in the last stage of our study, this possibility was investigated in depth. In particular, as at the industrial level the sorting machines require the use of few wavelengths, a variable selection procedure to identify the most effective wavelengths, or better some narrow intervals, to be used for model building was carried out. As described in Section 3.3 a procedure based on the application of the simple but effective iPLS-DA was adopted. The forward iPLS-DA algorithm for variable selection led to the selection of the wavelength intervals shown in Fig. 10.

The figure shows the RMSECV obtained for each interval (with the average spectrum of all the selected pixels used as training set superimposed as a black line). The numerical values inside the axes along the bottom (just above the wavelengths) indicate the number of latent variables (LVs) used to obtain the given RMSECV. The three green intervals are the selected intervals. The horizontal red dashed line indicates the RMSECV obtained when using all variables and 5 LVs. 3 intervals of 4 variables were selected, corresponding to the wavelengths between 1209 and 1230 nm, 1489–1510 nm and 1601–1622 nm.

The PLS-DA model based only on the 3 selected intervals obtained by iPLS-DA was then applied to the training image (wheat in bulk) and to the image with wheat grains in parallel lines. The results are reported in Figs. 8b and 9c, in terms of prediction maps. It is also evident from both the figures that accurate predictions were obtained for all the three classes and that the few misclassifications observed were related to pixels which are at the boundary between the kernels and the masked background.

4.5. *Comparison of the classification results*

Comparing the classification results obtained using 92 and 12 wavelengths, it appears that the adopted PLS-DA model allows to correctly classify the different classes of durum wheat, in both cases. Such results can be evaluated taking into account the values of sensitivity and specificity obtained for both classification models. Sensitivity is defined as the proportion of class members correctly classified; while specificity refers to the proportion of non-class members correctly classified. These parameters range from 0 to 1 with 1 the ideal value for a prediction model. Good results have been obtained as both sensitivity and specificity are between 0.94 and 0.99 in calibration and between 0.91 and 0.99 in cross-validation for the 92 wavelength model; for the 12 wavelengths model sensitivity and specificity range in calibration between 0.93 and 1.00 and in cross-validation between 0.92 and 1.00. The classification error was 0.01 (vitreous) and 0.04 (both

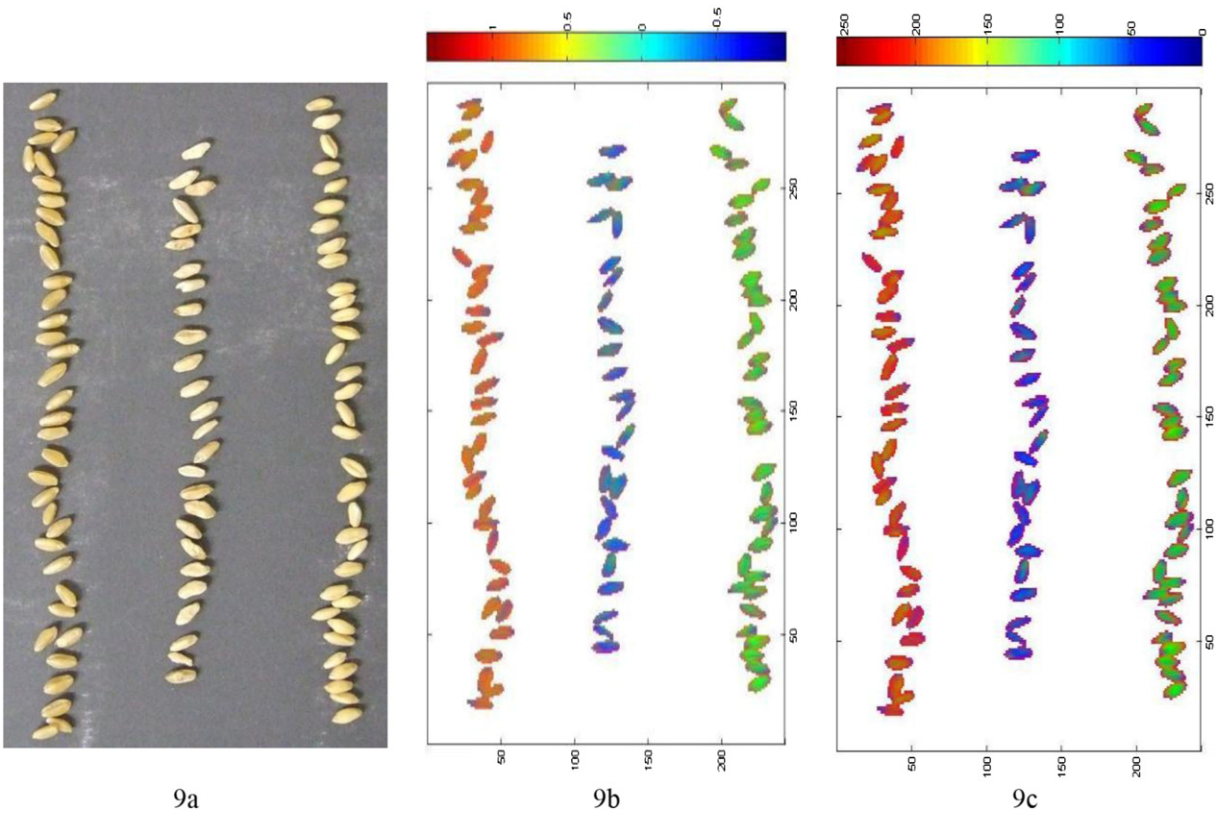


Fig. 9 – Source (9a) and prediction images based on PLS-DA model built for the classification of wheat in lines using 92 wavelengths (9b) and 12 wavelengths, respectively (9c) (red: vitreous, blue: *Fusarium*-damaged and green: yellow berry).

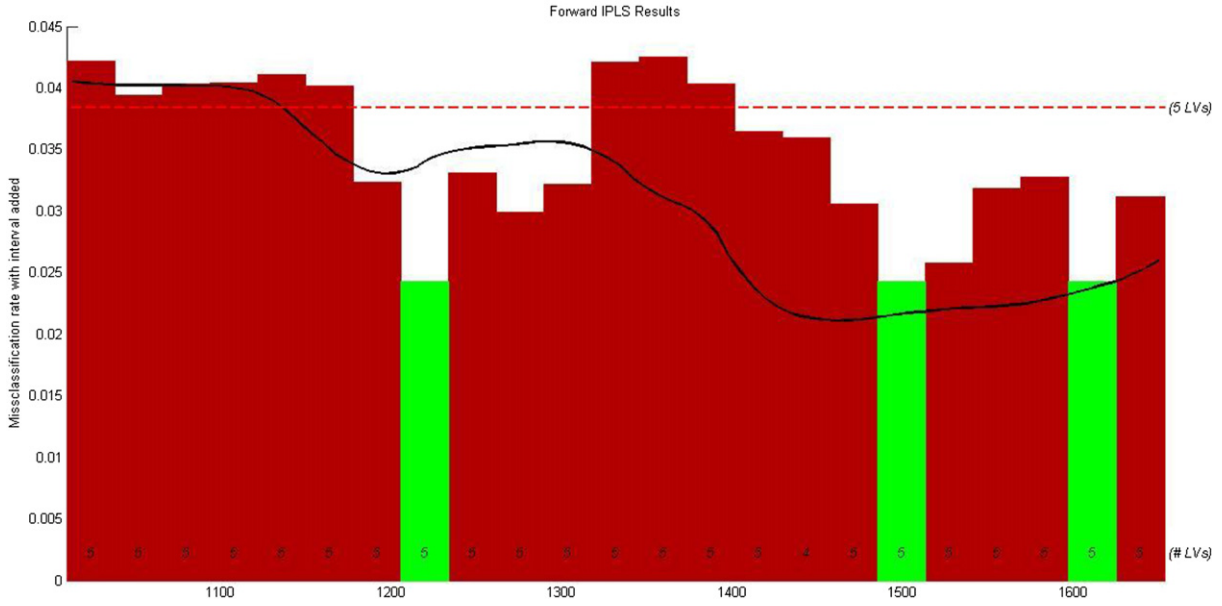


Fig. 10 – Selected intervals by applying the forward iPLS-DA. The *y*-axis represents the values of RMSECV (range: 0–1), the *x*-axis represents the wavelength range (nm). The black line represents the average reflectance spectrum of all the selected pixels used as training set. The dashed red line shows the RMSECV values obtained using all variables and 5 latent variables. The green columns indicate the selected three effective intervals of 4 wavelengths each. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

yellow berry and *Fusarium*-damaged) in the first model and 0.01 (vitreous) and 0.06 (both yellow berry and *Fusarium*-damaged) in the second. Such values are good and demonstrate that not only the model based on the entire investigated wavelength range, but also the one based on the selected 12 "effective" wavelengths, have a strong discriminant power in terms of wheat characteristic attribute recognition.

Comparing the results obtained for the different durum wheat kernel types, the best classification was obtained for vitreous class, with very low misclassification error (<1%), the yellow berry class showed slightly better results than *Fusarium*-damaged class, although in both cases the misclassification error was low, as previously stated.

5. Conclusions

An HSI system in the NIR region was investigated to evaluate the possibility of an objective, fast and non-destructive method to identify different classes of wheat, with particular reference to vitreous, *Fusarium*-damaged and yellow berry kernels. Conventional methods based on optical filters present some weaknesses, due to the very similar reflectance characteristics of the durum wheat typologies in the visible wavelength range. The NIR hyperspectral data were processed with multivariate statistical analysis methods, such as PCA, PLS-DA and iPLS-DA, to reduce their spectral dimension and redundancy and to extract useful image features for differentiating the three wheat classes. The results demonstrated that the classification is good, both using the full range of wavelengths (92 wavelengths from 1013 to 1650 nm), and selecting, through the described procedure, "only" 3 windows of 4 wavelengths each (1209–1230 nm, 1489–1510 nm and 1601–1622 nm). Following this approach it is therefore possible to greatly reduce the time needed to handle the hyperspectral data. As a result is envisaged the possibility to perform a real time monitoring of the process. The proposed approach could thus be profitably utilised for classification purposes and clearly shows the potentiality of NIR hyperspectral imaging to screen grain samples according to their spectral information. To reach this "*real time*" goal further tests will be carried out. The adoption and implementation of such an approach could be utilised for rapid quality control purposes and/or to realise a sorting of the products.

Authoring

Dr. Silvia Serranti conceived the work, defined the experimental set up and the procedures to apply. Eng. Daniela Cesare applied the chemometrics based procedures to the hyperspectral data and, together with Dr. Silvia Serranti, analysed the results and wrote text. Prof. Giuseppe Bonifazi contributed to results evaluation and paper final writing.

Acknowledgements

Thanks are given to Mr. Antonio and Andrea Uzzo of SEA s.r.l. (Imola, BO, Italy) for having provided the wheat samples.

REFERENCES

Bacci, L., Rapi, B., Colucci, F., & Novaro, P. (2002). Durum wheat quality evaluation software. In *Proceedings of the World Congress of Computers in Agriculture and Natural Resources* (pp. 49–55). St. Joseph, Mich: ASAE.

Baker, J. E., Dowell, F. E., & Throne, J. E. (1999). Detection of parasitized rice weevils in wheat kernels with near-infrared spectroscopy. *Biological Control, 16*, 88–90.

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics, 17*, 166–173.

Blanco, F., López-Mesas, M., Serranti, S., Bonifazi, G., Havel, J., & Valiente, M. (2012). Hyperspectral imaging based method for fast characterization of kidney stone types. *Journal of Biomedical Optics, 17*(7), 076027-1–076027-12.

Bonifazi, G., Damiani, L., Serranti, S., Bakker, E. J., & Rem, P. C. (2009). Innovative sensing technologies applied to post-consumer polyolefins recovery. *Metalurgia International, 14*, 5–10.

Bonifazi, G., & Serranti, S. (2006). Imaging spectroscopy based strategies for ceramic glass contaminants removal in glass recycling. *Waste Management, 26*, 627–639.

Bonifazi, G., Serranti, S., Bonoli, A., & Dall'Ara, A. (2009). Innovative recognition-sorting procedures applied to solid waste: the hyperspectral approach. In C. A. Brebbia, M. Neophytou, E. Beriatos, I. Ioannou, & A. G. Kungolos (Eds.), *Sustainable development and planning IV. Book series: WIT transactions on ecology and the environment, Vol. 120* (pp. 885–894). Southampton, Boston: WIT Press, ISBN 978-1-84564-422-2.

Davies, E. R. (2000). *Image processing for the food industry*. Singapore: World Scientific.

Del Fiore, A., Reverberi, M., Ricelli, A., Pinzari, F., Serranti, S., Fabbri, A. A., et al. (2010). Early detection of toxigenic fungi on maize by hyperspectral imaging analysis. *International Journal of Food Microbiology, 144*, 64–71.

Dowell, F. E. (2000). Differentiating vitreous and nonvitreous durum wheat kernels by using near-infrared spectroscopy. *Cereal Chemistry, 77*(2), 155–158.

Dowell, F. E., Throne, J. E., & Baker, J. E. (1998). Automated nondestructive detection of internal insect infestation of wheat kernels by using near-infrared spectroscopy. *Journal of Economic Entomology, 91*, 899–904.

Dowell, F. E., Throne, J. E., Wang, D., & Baker, J. E. (1999). Identifying stored-grain insects using near-infrared spectroscopy. *Journal of Economic Entomology, 92*, 165–169.

Fortunato de Carvalho Rocha, W., Post Sabin, G., Março, P. H., & Poppi, R. J. (2011). Quantitative analysis of piroxicam polymorphs pharmaceutical mixtures by hyperspectral imaging and chemometrics. *Chemometrics and Intelligent Laboratory Systems, 106*(2), 198–204.

Geladi, P., Grahn, H., & Burger, J. (2007). Techniques and applications of hyperspectral image analysis. In H. Grahn, & P. Geladi (Eds.), *Multivariate images, hyperspectral imaging: Background and equipment* (pp. 1–15), West Sussex, England.

Gosselin, R., Rodrigue, D., & Duchesne, C. (2011). A hyperspectral imaging sensor for on-line quality control of extruded polymer composite products. *Computers and Chemical Engineering, 35*, 296–306.

Gowen, A. A., O'Donnell, C. P., Cullen, P. J., & Bellc, S. E. J. (2008). Recent applications of chemical imaging to pharmaceutical process monitoring and quality control. *European Journal of Pharmaceutics and Biopharmaceutics, 69*, 10–22.

Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology, 18*, 590–598.

Haff, R. P., & Slaughter, D. C. (2004). Real-time X-rays inspection of wheat for infestation by the granary weevil *Sitophilus granarius* (L.). *Transactions of the ASAE, 47*, 531–537.

Hyvarinen, T., Herrala, E., & Dall'Ava, A. (1998). Direct sight imaging spectrograph: a unique add-on component brings spectral imaging to industrial applications. In. *Proceedings of SPIE, Vol. 3302*.

Jolivot, R., Vabres, P., & Marzani, F. (2011). Reconstruction of hyperspectral cutaneous data from an artificial neural network-based multispectral imaging system. *Computerized Medical Imaging and Graphics, 5*, 85–88.

Karunakaran, C., Jayas, D. S., & White, N. D. G. (2003). X-ray image analysis to detect infestations caused by insects in grain. *Cereal Chemistry, 80*, 553–557.

Kubik, M. (2007). Physical techniques in the study of art, archaeology and cultural heritage. In D. Creagh, & D. Bradley (Eds.), *Hyperspectral imaging: A new technique for the non-invasive study of artworks* (pp. 199–259). New York: Elsevier.

Leitner, R., Mairer, H., & Kercek, A. (2003). Real-time classification of polymers with NIR spectral imaging and blob analysis. *Real-Time Imaging, 9*, 245–251.

Luo, X., Jayas, D. S., & Symons, S. J. (1999). Identification of damaged kernels in wheat using a color machine vision system. *Journal of Cereal Science, 30*(1), 49–59.

Maghirang, E. B., Dowell, F. E., Baker, J. E., & Throne, J. E. (2003). Automated detection of single wheat kernels containing live or dead insects using near-infrared reflectance spectroscopy. *Transactions of the ASAE, 46*, 1277–1282.

Mahesh, S., Manickavasagan, A., Jayas, D. S., Paliwal, J., & White, N. D. G. (2008). Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes. *Biosystems Engineering, 101*(1), 50–57.

Mares, D. J. (1993). Pre-harvest sprouting in wheat: I. Influence of cultivar, rainfall and temperature during grain ripening. *Australian Journal of Agricultural Research, 44*, 1259–1272.

Martens, H., Høy, M., Wise, B. M., Bro, R., & Brockhoff, P. B. (2003). Pre-whitening of data by covariance-weighted pre-processing. *Journal of Chemometrics, 17*(3), 153–165.

Moller, B., & Munch, L. (2009). Infrared spectroscopy for food quality analysis and control. In D. W. Sun (Ed.), *Cereals and cereal products* (pp. 275–319). USA: Academic Press.

Neethirajan, S., Jayas, D. S., & White, N. D. G. (2007). Detection of sprouted wheat kernels using soft X-ray image analysis. *Journal of Food Engineering, 81*, 509–513.

Nørgaard, L., Saudaland, A., Wagner, J., Nielsen, J. P., Munck, L., & Engelsen, S. B. (2000). Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy. *Applied Spectroscopy, 54*, 413–419.

Otto, M. (1999). *Chemometrics, statistics and computer application in analytical chemistry*. New York: Wiley-VCH.

Ruan, R., Ning, S., Luo, L., Chen, P., Jones, R., Wilcke, W., et al. (2001). Estimation of weight percentage of scabby wheat kernels using an automatic machine vision and neural network based system. *Transactions of the ASAE, 44*(4), 983–988.

Sapirstein, H. D., & Bushuk, W. (1989). Quantitative determination of foreign material and vitreousity in wheat by digital image analysis. In H. Salovaara (Ed.), *Proceedings of ICC symposium, 89th. Wheat end-use properties* (pp. 453–474). Helsinki, Finland: University of Helsinki.

Sardana, V. (2000). Grain quality of wheat as influenced by application of nitrogen, climate and cultural practices: a review. *Agricultural Reviews, 21*(2), 116–120.

Serranti, S., Gargiulo, A., & Bonifazi, G. (2011). Characterization of post-consumer polyolefin wastes by hyperspectral imaging for quality control in recycling processes. *Waste Management, 31*, 2217–2227.

Serranti, S., Gargiulo, A., & Bonifazi, G. (2012). Classification of polyolefins from building and construction waste using NIR hyperspectral imaging system. *Resources Conservation & Recycling, 61*, 52–58.

Shahin, M. A., Dorrian, E., Symons, S. J. (2005). Proceedings of the CSAE/SCGR annual conference, Winnipeg, June. Machine vision system to detect hard vitreous kernels in durum wheat.

Shahin, M. A., & Symons, S. J. (2011). Detection of *Fusarium* damaged kernels in Canada Western Red Spring wheat using visible/near-infrared hyperspectral imaging and principal component analysis. *Computers and Electronics in Agriculture, 75*, 107–112.

Shatadal, P., Symons, S. J., & Dexter, J. E. (1998). Detecting hard vitreous kernels in durum wheat using image histogram. ASAE Paper No. 986028 *Proceedings of the international meeting of ASAE. St. Joseph, Michigan*.

Singh, C. B., Paliwal, J., Jayas, D. S., & White, N. D. G. (2006). *Near-infrared spectroscopy: Applications in the grain industry*. Winnipeg: CSBE, Paper no. 06-189.

Sugiyama, J. (1999). Visualization of sugar content in the flesh of a melon by near-infrared imaging. *Journal of Agricultural and Food Chemistry, 47*, 2715–2718.

Sun, D. W. (2010). *Hyperspectral imaging for food quality analysis and control*. San Diego CA: Academic Press/Elsevier.

Symons, S. J., Van Schepdael, L., & Dexter, J. E. (2003). Measurement of hard vitreous kernels in durum wheat by machine vision. *Cereal Chemistry, 80*(5), 511–517.

Tatzer, P., Wolf, M., & Panner, T. (2005). Industrial application for inline material sorting using hyperspectral imaging in the NIR range. *Real-Time Imaging, 11*, 99–107.

Wang, N., Zhang, N., Dowell, F. E., & Pearson, T. (2005). Determining vitreousness of durum wheat using transmitted and reflected images. *Transactions of the ASAE, 48*(1), 219–222.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems, 2*, 37–52.

Zayas, I. Y., Bechtel, D. B., Wilson, J. D., & Dempster, R. E. (1994). Distinguishing selected hard and soft red winter wheat by image analysis of starch granules. *Cereal Chemistry, 71*(1), 82–88.